B6

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

MR 43

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12P 33/02, C07K 3/00 A61K 37/24

A1

(11) International Publication Number: WO 92/14834

(43) International Publication Date: 3 September 1992 (03.09.92)

(21) International Application Number: PCT/US92/01196

(22) International Filing Date: 13 February 1992 (13.02.92)

(30) Priority data: 658,410 14 February 1991 (14.02.91) US

(71) Applicant: THE WHITTIER INSTITUTE FOR DIABETES AND ENDOCRINOLOGY [US/US]; 9894 Genesee Avenue, La Jolla, CA 92037 (US).

(72) Inventors: SHIMASKI, Shunichi ; 4427 Cather Avenue, San Diego, CA 92122 (US). LING, Nicholas, Chai-Kwan ; 5324 Bloch Street, San Diego, CA 92122 (US).

(74) Agents: WATT, Phillip, H. et al.; Room 900, 135 South LaSalle Street, Chicago, IL 60603 (US).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), MC (European patent), NL (European patent), SE (European patent).

Published
With international search report.

(54) Title: INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN

(57) Abstract

Two insulin-like growth factor binding proteins are isolated from rat serum, purified to homogeneity and partially sequenced. Using nucleotide probes based upon the amino terminal sequence of one of the isolated proteins, the complete sequence for the mature 252-residue rat protein, termed IGFBP-5, is deduced. The highly homologous 252-residue sequence of the human protein is thereafter separately deduced. These three proteins are useful in the inhibition of cell differentiation and/or proliferation requiring IGFs and are particularly useful in combating breast and bone cancers. Antibodies to the proteins may be employed in diagnostic assays, in purification of the protein and in the modulation of bone growth.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	RO	Romania
CA	Canada	IT	Italy	RU	Russian Federation
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		
ES	Spain	MG	Madagascar		

INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN

This invention relates to controlling the effects of insulin-like growth factors (IGFs) in mammals and more particularly to novel insulin-like growth factor binding proteins which can be employed to complex IGFs and thereby modulate IGF actions.

Background of the Invention

Two insulin-like growth factors (IGF-I and IGF-II) are presently known to exist and to be required for the proliferation of various cells. For example, the topical use of IGF-II for wound-healing is taught in U.S. Patent No. 4,885,163 (December 5, 1989). It is also reported that IGFs have a particular effect upon the growth of cells of mesodermal origin and on their differentiation, and further, that the IGFs exhibit potency in stimulating DNA synthesis in human fibroblasts and in rat osteoblasts. In addition, it is suggested that IGF-I may serve to stimulate collagen synthesis in human fibroblasts, whereas studies report that IGF-II may have a predominant role in undifferentiated cell proliferation.

Several proteins have been discovered which bind to these IGFs and modulate IGF actions either in an inhibitory or a stimulatory manner, and these proteins are termed insulin-like growth factor binding proteins (IGFBPs).

Insulin-like growth factors (IGF-I and IGF-II) are synthesized by multiple tissues and circulate in plasma to modulate the growth of various cell types. They do not exist in the blood as free hormones but are bound to carriers in the form of IGFBPs. To date four distinct classes of IGFBPs have been characterized, based on their complete primary structure having been obtained by molecular cloning, all of which are able to bind both

IGF-I and IGF-II and to modulate IGF actions either in an inhibitory or a stimulatory manner.

Based on the recommendation proposed by an IGFBP conference held in Vancouver, Canada in June 1989, the first BP class whose complete primary structure was deduced has been named IGFBP-1; its structure was deduced from cDNA clones identified in the libraries prepared from a human HEP-G2 hepatoma cell line, from human placenta and from both human and rat decidua. A human genomic clone encoding IGFBP-1 was also isolated and characterized (Brinkman, A., et al., B.B.R.C. 157, 898-907 (1989)), the gene locus of which is mapped at location p12-p13 on chromosome 7, Alitalo, T. et al., Hum. Genet. 83, 335-338 (1989). This protein exhibits a molecular weight (M_r) of 28-30 kDa on SDS/PAGE under non-reducing conditions and has almost equal binding affinity for IGF-I and IGF-II. It contains no potential N-linked glycosylation sites, but it has at least five potential O-linked glycosylation sites, which may account for a reported 4.3% carbohydrate content of the protein. The circulating level of the IGFBP-1 is elevated in patients and animals with insulin-dependent diabetes mellitus.

The second BP class is one for which the complete primary structure was deduced from cDNAs isolated from a rat BRL-3A cell library as well as from adult rat liver and human fetal liver libraries; it has been named IGFBP-2. This BP, having M_r of about 33-35 kDa on SDS/PAGE under non-reducing conditions, exhibits equal affinity for IGF-I and IGF-II when IGF-I is used as a radioligand, but it shows a marked preference for IGF-II when the radioligand is IGF-II. The level of IGFBP-2 in rat serum is high in fetus but decreases in adult. The physiological role of IGFBP-2 is not well known, but Adashi et al. Endocrinology, 126, 1305-1307 (1990) recently reported that pituitary follicle-stimulating hormone (FSH) inhibits the constitutive release of

IGFBP-2 from rat ovarian granulosa cells. The gene encoding the human protein has been mapped to chromosome 2, and the corresponding rat gene has also now been determined to be located on chromosome 2.

The third BP class is a high molecular weight IGFBP within the 150 kDa IGF-binding complex found in plasma. Its complete primary structure was deduced from human, porcine and rat cDNAs, and it has been named IGFBP-3, see Shimasaki, S., et al., B.B.R.C., 165, 907-912 (1989). This 150 kDa complex consists of three components, an IGFBP-3 of 53 kDa bound to an IGF and an acid-labile 80 kDa protein which can only bind to IGFBP-3 in association with IGF under neutral conditions. Both the 53 kDa IGFBP-3 and the 80 kDa acid-labile subunits are glycosylated. Moreover, the circulating level of the complex is dependent on growth hormone (GH). This protein has recently been isolated from ovarian follicular fluid, and it appears to act as an inhibitor to the FSH-stimulated production of estradiol in cultures of rat ovarian granulosa cells.

The fourth class of BP was isolated from human bone cell-conditioned medium, Mohan et al., P.N.A.S. 86, 8338-8342 (1989) and adult rat serum, Shimonaka et al., B.B.R.C. 165, 189-195 (1989). Its complete primary structure was subsequently deduced from a rat liver and a human placenta library and named IGFBP-4, see Shimasaki et al., Mol. Endocrinol. 4, 1451-58 (1990). Unlike the other three IGFBPs, this BP contains two extra cysteines in the midportion of the molecule in addition to the 18 homologous cysteines found in the other BPs. Moreover, it contains one potential Asn-linked glycosylation site.

Besides these four IGFBPs, it is believed that other BPs for IGF exist.

Summary of the Invention

Two additional novel IGFBPs have been isolated from rat serum, purified to homogeneity and partially

sequenced beginning at the N-terminus of each. Subsequently, the cDNAs encoding the complete primary structure of one of these proteins have been isolated and characterized from both the rat and the human species. Homologous cDNAs from other mammalian species can be likewise obtained and amino acid sequences of the IGFBP deduced. By homologous is meant having at least about 80% identity on a nucleotide base level. The deduced amino acid sequences of the cDNAs reveal a mature polypeptide of 252 amino acids for both the rat protein and the human protein which are highly homologous, i.e. at least about 90% identity on an amino acid level, and they contain what are believed to be signal sequences of 19-20 amino acid residues. These rat and human proteins which are hereinafter referred to as IGFBP-5, bind to both IGF-I and IGF-II and can be administered as anti-neoplastic agents along with an appropriate pharmaceutically or veterinarially acceptable carrier for various therapeutic purposes, such as the inhibition of cell differentiation and/or proliferation requiring IGFs. For example, IGFBP-5 can be used to combat breast and bone cancers and other tissues having a high IGF requirement, and in addition, it is expected to be useful for modulating bone growth. They can also be used in affinity chromatography columns for the purification of IGF-I and IGF-II.

Detailed Description of the Preferred Embodiments

IGFBP-5 and the other protein, herein referred to as IGFBP-6, were isolated and purified from adult rat serum using techniques generally similar to those described in Shimonaka, M. et al., B.B.R.C., 165, 189-195 (1989). Following gel filtration of the serum protein on Sephacryl S-200 superfine in 30% acetic acid, the gel-filtered fractions containing the BPs were located by a binding assay using $[^{125}I]$IGF-I. Fractions containing the BPs were pooled and dialyzed in a 3,500 M_r cut-off

Spectra/Por membrane (Spectrum Medical Industries, Los Angeles, CA) against 10 liters PBS/NaN_3 buffer, consisting of 20 mM sodium phosphate, 130 mM sodium chloride, 0.02% sodium azide, pH 7.4. After dialysis, the retentate was applied onto an IGF-II-coupled Affi-Gel 15 column at 20 ml/hr through a peristaltic pump in a cold room. After all the sample had been pumped through the column, the gel bed was washed with 200 ml PBS buffer, containing 0.5 M NaCl, at the same flow rate. The adsorbed proteins were eluted with 0.5 M acetic acid, pH 3.0, at 3.5 ml/hr, and 1 ml fractions were collected. The IGFBPs in the eluate fractions were located by UV absorbance at 280 nm.

The recovered IGFBPs from the affinity column were pooled and, after dilution with an equal volume of water, were pumped directly onto an 0.7 x 25 cm Aquapore RP-300, 10 μm particle size, C_8 column (Applied Biosystems, Inc., Santa Clara, CA) at a flow rate of 3 ml/min. After loading, the adsorbed proteins were separated in a model 322 gradient HPLC system (Beckman, San Ramon, CA) using a linear gradient of 18-36% acetonitrile in the 0.1% trifluoroacetic acid (vol/vol) solvent system in 180 min at a flow rate of 3 ml/min, as described generally in Ui et al., Endocrinology 125, 912-916 (1989). The column effluent was monitored by UV absorbance at 210 nm. The chromatogram of this HPLC step showed a number of potential peaks, and each peak was further fractioned by another PHLC step on a 1 x 25 cm Vydac, 5 μm particle size, C_4 column (Separations Group, Hesperia, CA) using a linear gradient of 14-26% acetonitrile in the 0.1% triethylammonium phosphate (vol/vol) solvent system in 120 min at a flow rate of 1 ml/min. Each of the recovered HPLC peaks was subjected to microsequence analysis in an ABI model 470A gas-phase protein sequenator.

As a result of this sequencing of the amino termini of the purified proteins, the sequence of the

first 36 amino acid residues of IGFBP-5 was obtained, along with the first 13 residues of IGFBP-6. In this manner, it was determined that the amino terminus of the rat protein IGFBP-5 had the following sequence (SEQ ID NO:1):

Leu-Gly-Ser-Phe-Val-His-Xaa-Glu-Pro-Xaa-Asp-Glu-Lys-Ala-Leu-Ser-Met-Xaa-Pro-Pro-Ser-Pro-Leu-Gly-Xaa-Glu-Leu-Val-Lys-Glu-Pro-Gly-Xaa-Gly-Xaa-Xaa.

Based upon the evident homology between this protein and the other 4 known IGFBP structures, it was assumed that the unidentified amino acid residues (Xaa) were cysteine (Cys).

It was similarly determined that the amino terminus of the rat protein IGFBP-6 had the following sequence (SEQ ID NO:2):

Ala-Leu-Ala-Gly-Xaa-Pro-Gly-Xaa-Gly-Pro-Gly-Val-Gln.

Again, based upon the evident homology between this protein and the other IGFBP structures, it is assumed that the unidentified amino acid residues (Xaa) are cysteine.

A similar isolation and purification procedure was carried out using porcine follicular fluid (pFF) resulting in recovered IGFBPs including porcine IGFBP-6. By amino acid sequence analysis, the N-terminus of the homogeneous porcine IGFBP-6 was determined to be (SEQ ID NO:7):

Ala-Gln-Xaa-Pro-Gly-Xaa-Gly-Gln-Gly-Val-Gln-Thr-Gly-Xaa-Pro-Gly.

Again, based upon the evident homology between this protein and the other IGFBP structures, it is assumed

that the unidentified amino acid residues (Xaa) are cysteine.

The amino acids are referred to herein using either the standard 3-letter or 1-letter designations as follows:

NAME	3-LETTER	1-LETTER
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

To obtain the DNA sequence of a desired protein, a cDNA library or an expression library is often produced in a conventional manner by reverse transcription from messenger RNA (mRNA) from a mammalian cell line or tissue. To select clones containing DNA encoding the desired protein sequences, a hybridization probe obtained by PCR technology (or mixed probes which accommodat the degeneracy of the genetic code and correspond to a

selected portion of the target protein are produced) is used to hydridize with clones containing such sequences. Screening of such an expression library with antibodies to all or a portion of the protein may also be used, either alone or in conjunction with hybridization probing, to identify or confirm the presence of cDNA library clones which are expressing the target protein.

Using synthetic DNA encoding for the amino-terminal region of IGFBP-5, which is based upon all possible codon combinations, it was decided to utilize the polymerase chain reaction (PCR) technology to generate the native nucleotide sequence for this region. Accordingly, two synthetic oligonucleotide mixture primers for PCR were designed which incorporated all possible codon combinations encoding (a) a 7-amino acid residue sequence near the N-terminus selected from SEQ ID NO:1, namely, Phe-Val-His-Cys-Glu-Pro-Cys- and (b) the 7-amino acid residue sequence which appears at the carboxy-terminus of this 36-residue SEQ ID NO:1, namely Glu-Pro-Gly-Cys-Gly-Cys-Cys. The synthetic primer mixtures are as follows:

5'-TT(CT)GT(ACGT)CA(CT)TG(CT)GA(AG)CC(ACGT)TG-3' and
3'-CT(CT)GG(ACGT)CC(ACGT)AC(AG)CC(ACGT)AC(AG)AC-5'

PCR was performed by a TwinBlock™ system (Ericomp, San Diego, CA) with GeneAmp™ DNA Amplification Reagent Kit (Perkin-Elmer Cetus, Norwalk, CT) using a PMSG-stimulated rat ovary cDNA library as a template. Annealing reactions were performed at 60°C. for 30 sec., followed by a 30 sec. extension at 72°C. and 15 sec. denaturation at 94°C. After 35 cycles of amplification, a PCR-extended fragment of 98 bp was purified, kinased by ATP and then cloned into EcoRV site of pBluescript SK+ (Stratagene, San Diego, CA). The DNA sequence of the PCR-amplified cDNA fragment was determined by the double-

stranded dideoxy-chain-termination method using Sequenase (United States Biochem. Co., Cleveland, OH); its deduced amino acid sequence matched with that obtained by protein sequencing.

This PCR-amplified cDNA fragment having the expected 98 bp length obtained from the rat ovary cDNA library was thereafter labeled by a random priming method using ^{32}P-dCTP and used as a probe to screen the same rat ovary cDNA library to isolate cDNA clones encoding rat IGFBP-5. Upon screening with this probe, six positive clones were obtained from one million independent clones. Each of these clones was sequenced, and the results revealed that all six of them contained the complete coding region of rat IGFBP-5 including the signal sequence.

The complete DNA sequence of one clone is shown in TABLE 1 wherein an open reading frame encodes a protein of 271 amino acid residues with the predicted amino acid sequence being shown below each codon. The amino-terminal residue of the 252-residue mature protein is denoted by +1 so as to be in agreement with the amino-terminal residue of the purified rat IGFBP-5. The preceding 19 amino acid sequence leading to the amino-terminal residue of the mature protein fits a typical signal peptide sequence, terminating in a neutral residue with a small side-chain which, in this case, is Gly at position -1. The complete 271-residue sequence is set forth as follows as SEQ ID NO:3:

```
Met Val Ile Ser Val Val Leu Leu Leu Leu Ala Ala Cys Ala Val Pro
1               5                   10                  15

Ala Gln Gly Leu Gly Ser Phe Val His Cys Glu Pro Cys Asp Glu Lys
            20                  25                  30

Ala Leu Ser Met Cys Pro Pro Ser Pro Leu Gly Cys Glu Leu Val Lys
        35                  40                  45

Glu Pro Gly Cys Gly Cys Cys Met Thr Cys Ala Leu Ala Glu Gly Gln
    50                  55                  60
```

Ser	Cys	Gly	Val	Tyr	Thr	Glu	Arg	Cys	Ala	Gln	Gly	Leu	Arg	Cys	Leu
65					70					75					80
Pro	Arg	Gln	Asp	Glu	Glu	Lys	Pro	Leu	His	Ala	Leu	Leu	His	Gly	Arg
				85					90					95	
Gly	Val	Cys	Leu	Asn	Glu	Lys	Ser	Tyr	Gly	Glu	Gln	Thr	Lys	Ile	Glu
			100					105					110		
Arg	Asp	Ser	Arg	Glu	His	Glu	Glu	Pro	Thr	Thr	Ser	Glu	Met	Ala	Glu
		115					120					125			
Glu	Thr	Tyr	Ser	Pro	Lys	Val	Phe	Arg	Pro	Lys	His	Thr	Arg	Ile	Ser
	130					135					140				
Glu	Leu	Lys	Ala	Glu	Ala	Val	Lys	Lys	Asp	Arg	Arg	Lys	Lys	Leu	Thr
145					150					155					160
Gln	Ser	Lys	Phe	Val	Gly	Gly	Ala	Glu	Asn	Thr	Ala	His	Pro	Arg	Val
				165					170					175	
Ile	Pro	Ala	Pro	Glu	Met	Arg	Gln	Glu	Ser	Asp	Gln	Gly	Pro	Cys	Arg
			180					185					190		
Arg	His	Met	Glu	Ala	Ser	Leu	Gln	Glu	Phe	Lys	Ala	Ser	Pro	Arg	Met
		195					200					205			
Val	Pro	Arg	Ala	Val	Tyr	Leu	Pro	Asn	Cys	Asp	Arg	Lys	Gly	Phe	Tyr
	210					215					220				
Lys	Arg	Lys	Gln	Cys	Lys	Pro	Ser	Arg	Gly	Arg	Lys	Arg	Gly	Ile	Cys
225					230					235					240
Trp	Cys	Val	Asp	Lys	Tyr	Gly	Met	Lys	Leu	Pro	Gly	Met	Glu	Tyr	Val
				245					250					255	
Asp	Gly	Asp	Phe	Gln	Cys	His	Ala	Phe	Asp	Ser	Ser	Asn	Val	Glu	
			260					265					270		

TABLE 1 shows the nucleotide sequence (SEQ ID NO:6) and deduced amino acid sequence of rat IGFBP-5 as determined by sequencing a cDNA clone. The nucleotides are numbered at the right, and the amino acids, in one-letter code, are numbered throughout. The amino terminus of the mature protein, Leu, is denoted by +1. There appears to be no potential Asn-linked glycosylation site.

TABLE 1

```
CTTTTTTTTCCTCTCCCTTGATTTCAACATTTTCCCGATCTTCTTGGCGCCGCCGACGCC    60
TCTTACCTGTTCTGCGCAGCGCGCGAGCTGGCAGCTGAGAGAAGTGGGGGTGCGTTTTAG   120
GTTTTAAGCAAAGGCAAAAAAAAAAAATTAAGCCAAATCCATTTTTTTTCCTTCACCTCTC   180
CCCGTTTCAAGGCCTCCAAGATCATTATTTCTGTAGCTTTGGGGTGAGCGATTCTGTGTT   240
TCTCTTCATCACCCCTCCAATTCTGCCCCGATCCCGCCTGGGTCTCCACTCACTGCGTGC   300
ACCTGGCGCGCCTCTTTTTTTTTTCACCCCCAACCTGTTGCAAGTCTTTAATCCTTGCAA   360
TTGGGACTTGCGTGCAGGCACCTGAATCCTCCTTGCCTCATATTTTGCAAGTGTTTGGAG   420
GACAGCACCTGCTTTACCTGCAAGAGATATTTTTTAAAAAAAAAAAAATCTCCAGGCTCC   480
CTCTTGGCCCCTTTCTCCACACACTCTCGCTCTCCTGCCCCGCCCCGAGGTAAAGCCAGA   540
                                          -10
CTCGGAAAAAATGGTGATCAGCGTGGTCCTCCTGCTGCTGGCCGCCTGTGCCGTGCCGGC   600
           M  V  I  S  V  V  L  L  L  L  A  A  C  A  V  P  A
  -1 +1                             +10
TCAAGGCCTGGGCTCTTTCGTGCATTGTGAACCCTGCGACGAGAAAGCTCTGTCCATGTG   660
 Q  G  L  G  S  F  V  H  C  E  P  C  D  E  K  A  L  S  M  C
  +20                               +30
TCCCCCCAGCCCTCTGGGCTGTGAGCTGGTCAAAGAGCCCGGCTGTGGCTGCTGCATGAC   720
 P  P  S  P  L  G  C  E  L  V  K  E  P  G  C  G  C  C  M  T
  +40                               +50
TTGCGCCCTGGCGGAGGGACAGTCGTGTGGCGTCTACACTGAGCGCTGCGCCCAGGGTTT   780
 C  A  L  A  E  G  Q  S  C  G  V  Y  T  E  R  C  A  Q  G  L
  +60                               +70
GCGCTGTCTCCCCCGGCAGGATGAGGAGAAGCCGCTGCACGCCCTGCTGCACGGCCGCGG   840
 R  C  L  P  R  Q  D  E  E  K  P  L  H  A  L  L  H  G  R  G
  +80                               +90
GGTTTGCCTCAACGAAAAGAGCTACGGCGAGCAAACCAAGATAGAGAGAGACTCTCGGGA   900
 V  C  L  N  E  K  S  Y  G  E  Q  T  K  I  E  R  D  S  R  E
  +100                              +110
GCATGAGGAACCCACCACCTCCGAGATGGCTGAGGAGACCTACTCCCCGAAGGTCTTCCG   960
 H  E  E  P  T  T  S  E  M  A  E  E  T  Y  S  P  K  V  F  R
  +120                              +130
GCCCAAGCACACTCGCATTTCCGAGCTGAAGGCCGAGGCTGTGAAGAAGGATCGCAGAAA  1020
 P  K  H  T  R  I  S  E  L  K  A  E  A  V  K  K  D  R  R  K
  +140                              +150
GAAGCTGACCCAGTCTAAGTTTGTGGGGGGCGCGGAGAACACTGCCCACCCCCGAGTCAT  1080
 K  L  T  Q  S  K  F  V  G  G  A  E  N  T  A  H  P  R  V  I
  +160                              +170
CCCTGCACCTGAGATGAGACAGGAATCTGACCAAGGCCCCTGCCGCAGACACATGGAAGC  1140
 P  A  P  E  M  R  Q  E  S  D  Q  G  P  C  R  R  H  M  E  A
  +180                              +190
TTCCCTCCAGGAGTTCAAAGCCAGCCCACGCATGGTGCCCCGTGCCGTGTACCTGCCCAA  1200
 S  L  Q  E  F  K  A  S  P  R  M  V  P  R  A  V  Y  L  P  N
  +200                              +210
CTGTGACCGCAAAGGATTCTACAAGAGAAAGCAGTGCAAGCCTTCTCGTGGCCGCAAACG  1260
 C  D  R  K  G  F  Y  K  R  K  Q  C  K  P  S  R  G  R  K  R
  +220                              +230
TGGCATCTGCTGGTGTGTGGACAAGTATGGGATGAAGCTGCCGGGCATGGAGTACGTCGA  1320
 G  I  C  W  C  V  D  K  Y  G  M  K  L  P  G  M  E  Y  V  D
  +240                              +250
TGGGGACTTTCAGTGCCACGCCTTCGACAGCAGTAACGTTGAGTGACGCGTCCCCTCCCT  1380
 G  D  F  Q  C  H  A  F  D  S  S  N  V  E END

TCCTCCCCCTTTCCTACCCCCCAGCCCCAACTCCAGCCAGCGCCTCCCTCCACCCCAGGA  1440
CGTCACTCATTTCATCTCATTTAGGGAAATATATATACATATATATTTGAGGAAACTGA  1500
GGACCTCGGAATCTCTAGCAAGGGCTAAGGAGACACTCCCCATTCCCGACCCCGGAAACG  1560
TATTCCTATTTGAAGCAAGTTGAACGGACAGAGAAGGGAAGAAGAGAAGGGGCAAGAAGG  1620
AGCGAGGAAT
```

The ins rt fragment of this clone was excised and subsequently used as a probe to identify the corresponding human IGFBP-5 clone. A human placenta cDNA library in phages consisting of about 2.4 x 10^6 independent clones was screened with the rat IGFBP-5 probe, and 98 positive clones were obtained out of a half million independent clones. Positive clones were then randomly selected and purified, and the insert DNAs were prepared. One of the longest clones was subcloned into the EcoRI site of pBluescript SK+ for DNA sequence determination. The complete DNA sequence of this clone as well as its deduced amino acid sequence are shown in TABLE 2. The open reading frame of this clone encodes a protein of 272 amino acids which is believed to include a signal sequence of 20 residues based on its homology with the rat sequence. The complete 272-residue sequence is set forth as follows as SEQ ID NO:4:

```
Met Val Leu Leu Thr Ala Val Leu Leu Leu Leu Ala Ala Tyr Ala Gly
1               5                   10                  15

Pro Ala Gln Ser Leu Gly Ser Phe Val His Cys Glu Pro Cys Asp Glu
            20                  25                  30

Lys Ala Leu Ser Met Cys Pro Pro Ser Pro Leu Gly Cys Glu Leu Val
        35                  40                  45

Lys Glu Pro Gly Cys Gly Cys Cys Met Thr Cys Ala Leu Ala Glu Gly
    50                  55                  60

Gln Ser Cys Gly Val Tyr Thr Glu Arg Cys Ala Gln Gly Leu Arg Cys
65                  70                  75                  80

Leu Pro Arg Gln Asp Glu Glu Lys Pro Leu His Ala Leu Leu His Gly
                85                  90                  95

Arg Gly Val Cys Leu Asn Glu Lys Ser Tyr Arg Glu Gln Val Lys Ile
            100                 105                 110

Glu Arg Asp Ser Arg Glu His Glu Glu Pro Thr Thr Ser Glu Met Ala
        115                 120                 125

Glu Glu Thr Tyr Ser Pro Lys Ile Phe Arg Pro Lys His Thr Arg Ile
    130                 135                 140
```

Ser Glu Leu Lys Ala Glu Ala Val Lys Lys Asp Arg Arg Lys Lys Leu
145 150 155 160

Thr Gln Ser Lys Phe Val Gly Gly Ala Glu Asn Thr Ala His Pro Arg
165 170 175

Ile Ile Ser Ala Pro Glu Met Arg Gln Glu Ser Glu Gln Gly Pro Cys
180 185 190

Arg Arg His Met Glu Ala Ser Leu Gln Glu Leu Lys Ala Ser Pro Arg
195 200 205

Met Val Pro Arg Ala Val Tyr Leu Pro Asn Cys Asp Arg Lys Gly Phe
210 215 220

Tyr Lys Arg Lys Gln Cys Lys Pro Ser Arg Gly Arg Lys Arg Gly Ile
225 230 235 240

Cys Trp Cys Val Asp Lys Tyr Gly Met Lys Leu Pro Gly Met Glu Tyr
245 250 255

Val Asp Gly Asp Phe Gln Cys His Thr Phe Asp Ser Ser Asn Val Glu
260 265 270

TABLE 2 shows the nucleotide sequence (SEQ ID NO:5) and deduced amino acid sequence of human IGFBP-5 as determined by sequencing a cDNA clone. The nucleotides are numbered at the right, and the amino acids in one-letter code are numbered throughout, as in TABLE 1.

TABLE 2

```
                                                                      -20
 CCCTGCACTCTCGCTCTCCTGCCCCACCCCGAGGTAAAGGGGGCGACTAAGAGAAGATGG          60
                                                                       M
                          -10                                     -1  +1
 TGTTGCTCACCGCGGTCCTCCTGCTGCTGGCCGCCTATGCGGGGCCGGCCCAGAGCCTGG         120
 V  L  L  T  A  V  L  L  L  L  A  A  Y  A  G  P  A  Q  S  L
                          +10                                     +20
 GCTCCTTCGTGCACTGCGAGCCCTGCGACGAGAAAGCCCTCTCCATGTGCCCCCCCAGCC         180
 G  S  F  V  H  C  E  P  C  D  E  K  A  L  S  M  C  P  P  S
                          +30                                     +40
 CCCTGGGCTGCGAGCTGGTCAAGGAGCCGGGCTGCGGCTGCTGCATGACCTGCGCCCTGG         240
 P  L  G  C  E  L  V  K  E  P  G  C  G  C  C  M  T  C  A  L
                          +50                                     +60
 CCGAGGGGCAGTCGTGCGGCGTCTACACCGAGCGCTGCGCCCAGGGGCTGCGCTGCCTCC         300
 A  E  G  Q  S  C  G  V  Y  T  E  R  C  A  Q  G  L  R  C  L
                          +70                                     +80
 CCCGGCAGGACGAGGAGAAGCCGCTGCACGCCCTGCTGCACGGCCGCGGGGTTTGCCTCA         360
 P  R  Q  D  E  E  K  P  L  H  A  L  L  H  G  R  G  V  C  L
                          +90                                     +100
 ACGAAAAGAGCTACCGCGAGCAAGTCAAGATCGAGAGAGACTCCCGTGAGCACGAGGAGC         420
 N  E  K  S  Y  R  E  Q  V  K  I  E  R  D  S  R  E  H  E  E
                          +110                                    +120
 CCACCACCTCTGAGATGGCCGAGGAGACCTACTCCCCCAAGATCTTCCGGCCCAAACACA         480
 P  T  T  S  E  M  A  E  E  T  Y  S  P  K  I  F  R  P  K  H
                          +130                                    +140
 CCCGCATCTCCGAGCTGAAGGCTGAAGCAGTGAAGAAGGACCGCAGAAAGAAGCTGACCC         540
 T  R  I  S  E  L  K  A  E  A  V  K  K  D  R  R  K  K  L  T
                          +150                                    +160
 AGTCCAAGTTTGTCGGGGGAGCCGAGAACACTGCCCACCCCCGGATCATCTCTGCACCTG         600
 Q  S  K  F  V  G  G  A  E  N  T  A  H  P  R  I  I  S  A  P
                          +170                                    +180
 AGATGAGACAGGAGTCTGAGCAGGGCCCCTGCCGCAGACACATGGAGGCTTCCCTGCAGG         660
 E  M  R  Q  E  S  E  Q  G  P  C  R  R  H  M  E  A  S  L  Q
                          +190                                    +200
 AGCTCAAAGCCAGCCCACGCATGGTGCCCCGTGCTGTGTACCTGCCCAATTGTGACCGCA         720
 E  L  K  A  S  P  R  M  V  P  R  A  V  Y  L  P  N  C  D  R
                          +210                                    +220
 AAGGATTCTACAAGAGAAAGCAGTGCAAACCTTCCCGTGGCCGCAAGCGTGGCATCTGCT         780
 K  G  F  Y  K  R  K  Q  C  K  P  S  R  G  R  K  R  G  I  C
                          +230                                    +240
 GGTGCGTGGACAAGTACGGGATGAAGCTGCCAGGCATGGAGTACGTTGACGGGGACTTTC         840
 W  C  V  D  K  Y  G  M  K  L  P  G  M  E  Y  V  D  G  D  F
                          +250
 AGTGCCACACCTTCGACAGCAGCAACGTTGAGTGATGCGTCCCCCCCCAACCTTTCCCTC         900
 Q  C  H  T  F  D  S  S  N  V  E END

 ACCCCCTCCCACCCCCAGCCCCGACTCCAGCCAGCGCCTCCCTCCACCCCAGGACGCCAC         960
 TCATTTCATCTCATTTAAGGAAAAATATATATCTATCTATTTGAAAAAAAAAAAAAAAAA        1020
 CCC
```

Th amino acid sequence comparison betwe n rat and human IGFBP-5 is shown in TABLE 3 wher in amino acids are shown in one-letter code, and only amino acids that differ from rat IGFBP-5 are presented in the human structure. The mature forms of rat IGFBP-5 and human IGFBP-5 consist of 252 amino acids, whereas the signal sequence of the human homolog is 1 amino acid longer.

TABLE 3

Amino acid sequences of rat and human IGFBP-5

```
                                         1    10        20        30        40        50        60        70
Rat   IGFBP-5   MVISV-VLLLLAACAVPAQGLGSFVHCEPCDEKALSMCPPSPLGCELVKEPGCGCCMTCALAEGQSCGVYTERCAQGLRCLPRQDEEKPLHAL
Human IGFBP-5   ..LLTA.......Y.G...S.........................................................................

                    80        90       100       110       120       130       140       150       160
Rat   IGFBP-5   LHGRGVCLNEKSYGEQTKIERDSREHEEPTTSEMAEETYSPKVFRPKHTRISELKAEAVKKDRRKKLTQSKFVGGAENTAHPRVIPAPEMRQE
Human IGFBP-5   .............R..V.........................I........................................I.S.......

                170       180       190       200       210       220       230       240       250
Rat   IGFBP-5   SDQGPCRRHMEASLQEFKASPRMVPRAVYLPNCDRKGFYKRKQCKPSRGRKRGICWCVDKYGMKLPGMEYVDGDFQCHAFDSSNVE
Human IGFBP-5   .E..............L..............................................................T.......
```

The proteins are highly homologous, about 97%, with most of the amino acid substitutions between the two species being located in the middle of the molecule. One gap at position -15 in the signal sequence is inserted in the rat sequence to allow maximal homology alignment. The location of all the cysteines are conserved, and no potential N-linked glycosylation site appears in either molecule.

When initial attempts to prepare a homologous probe using PCR and the N-terminal sequence information obtained with respect to the protein IGFBP-6 were not immediately successful, it was decided to generate additional amino acid sequence information. Accordingly, the homogeneous protein porcine IGFBP-6 was reduced by treatment with dithiothreitol and the free sulfhydryl groups were labeled with [^{14}C]-iodoacetamide. This labeled protein was digested with trypsin using a protein/enzyme ratio of 20:1 in 100 microliters 0.5M TRIS-HCl, pH 8.1. The digested peptide fragments were isolated by HPLC, and the determination of the amino acid sequences thereof was carried out using an ABI 470A protein sequenator. The results of this analysis produced the following 3 additional amino acid sequences from the homogeneous porcine IGFBP-6 protein:

(SEQ ID NO:8) Ala-Gly-Pro-Cys-Trp-Cys-Val-Asp-Ser-Arg-Pro-Asn-Pro-Gly-Gly-Val-Gln-Asp-Thr-Glu-Met-Gly-Pro-Cys-Arg;

(SEQ ID NO:9) Leu-Ala-Gln-Cys-Pro-Gly-Cys-Gly-Gln-Gly-Val-Gln-Thr-Gly-Cys-Pro-Gly-Gly-Cys-Ala-Glu-Glu-Glu-Asp-Gly-Gly-Xaa-Pro-Arg-Glu-Arg-Val;

(SEQ ID NO:10) Glu-Gly-Gln-Gln-Cys-Gly-Val-Tyr-Thr-Pro-Asn-Cys-Ala-Pro-Gly-Leu-Gln-Cys-Gln-Pro-Pro-Glu-Glu-Asp-Gln-Ala-Pro-Leu-Arg.

SEQ ID NO:10 was chosen for use to prepare DNA probes using PCR and using pig genomic DNA as a template. A PCR-amplified DNA fragment was obtained and was labeled as set forth hereinbefore and used to screen the PMSG-primed rat ovary cDNA library and also the human placenta library. Positive clones were located from both libraries.

Antibodies to these IGFBP-5 proteins of either monoclonal or polyclonal form can be produced using techniques presently known in the art, and antibodies which are effective to counteract the effects of IGFBP-5 can be elicited using only the synthetic N-terminal segment of the rat protein. For example, antibodies raised in rabbits against a synthetic peptide, representing the amino terminal sequence of the IGFBP-5, should recognize the synthetic peptide and the IGFBP-5 on an equimolar basis, and they should be capable of inhibiting the activity of the native protein in vitro. Amino terminal-directed antibodies to IGFBP-5 may be obtained, for example, by immunizing three month old male and female white New Zealand rabbits with the synthetic peptide to which Tyr has been added at the C-terminus in order to couple it, as an antigen, to BSA by a bisdiazotized benzidine(BDB) linkage by reaction for 2 hours at 4°C. The reaction mixture is dialyzed to remove low molecular weight material, and the retentate is frozen in liquid nitrogen and stored at -20°C. Animals are immunized with the equivalent of 1 mg of the peptide antigen according to the procedure of Benoit et al. P.N.A.S. USA, 79, 917-921 (1982). At four week intervals, the animals are boosted by injections of 200 μg of the antigen and bled ten to fourteen days later. After the third boost, examination of the antiserum should show its capacity to bind radioiodinated antigen peptide prepared by the chloramine-T method, and

then it would be purified by CMC-ion exchange column chromatography.

A radioimmunoassay can then be established with the antisera and serum from subsequent bleeds from the same rabbits. The native protein should be recognized by the antibodies on an equimolar basis as compared to the synthetic peptide antigen. These antibodies are considered to be capable of at least partially neutralizing the biological activity of the IGFBP-5, and substantially all such activity can likely be neutralized when higher amounts of antibodies are used. Immunoaffinity or affinity chromatography can also be applied to achieve the purification of IGFBP-5 from serum or biological materials; likewise IGFBP-5 can be used in affinity chromatography to purify IGF-I or IGF-II.

Antibodies to IGFBP-5 can be used in assays for detecting the levels of IGFBP-5 in mammals, particularly humans. The antibodies can also be used for treatment to neutralize the effect of IGFBP-5 in mammals and are useful for diagnostic test kits and the like.

From presently available evidence, it is most likely that there is internal disulfide-bonding between cysteine residues of the chain. Mammalian IGFBP-5 polypeptides produced by recombinant DNA techniques are considered to be inherently biologically active, and the three-dimensional structure which the IGFBP-5 assumes within cells is likely the structure recognized by the receptor. The three-dimensional structure which the molecule assumes through natural folding and through hydrophobic and hydrophilic interactions with aqueous media may also promote desired bonding or non-bonding between cysteine residues. Also, enzymatic regulatory mechanisms within cells may help to ensure desired disulfide bonding or non-bonding, either by preventing bonding or by directing disulfide bonding between particular cysteine residues. Enzymes might also cleave

"incorrect" bonding to enable the molecule to reorientate itself and assume the correct natural structure. Cysteine residues that are not internally bonded may be disulfide-bonded to free cysteine moieties. The three-dimensional structure of the molecule may also be such that random bonding or non-bonding of cysteine residues, either with each other or to free cysteines, does not substantially affect the biological structure of the protein molecule.

To synthesize a protein having the mammalian IGFBP-5 amino acid residue sequence, e.g. human IGFBP-5 (SEQ ID NO:4) by recombinant DNA, a double-stranded DNA chain which encodes IGFBP-5 might be synthetically constructed. Although it is nowadays felt that PCR techniques would be method of choice to produce DNA chains, a DNA chain encoding IGFBP-5 could be designed using certain particular codons that are more efficient for polypeptide expression in a certain type of organism, i.e. selection might employ those codons which are most efficient for expression in the type of organism which is to serve as the host for the recombinant vector. However, any correct set of codons will encode a desired product, although perhaps slightly less efficiently. Codon selection may also depend upon vector construction considerations; for example, it may be necessary to avoid placing a particular restriction site in the DNA chain if, subsequent to inserting the synthetic DNA chain, the vector is to be manipulated using the restriction enzyme that cleaves at such a site. Also, one should avoid placing restriction sites in the DNA chain if the host organism, which is to be transformed with the recombinant vector containing the DNA chain, is known to produce a restriction enzyme that would cleave at such a site within the DNA chain.

For example, a synthetic IGFBP-5-encoding DNA chain could be assembled by constructing oligonucleotides

by conventional procedures, such as thos described in T. Maniatis et al., Cold Spring Harbor Laboratory Manual, Cold Spring Harbor, New York (1982) (hereinafter, CSHLM). Sense and antisense oligonucleotide chains, up to about 70 nucleotide residues long, are synthesized, preferably on automated synthesizers, such as the Applied Biosystem Inc. Model 380A DNA synthesizer. The oligonucleotide chains are usually constructed so that portions of the sense and antisense oligonucleotides overlap, associating with each other through hydrogen bonding between complementary base pairs and thereby forming double stranded chains, in most cases with gaps in the strands. Subsequently, the gaps in the strands are filled in, and oligonucleotides of each strand are joined end to end with nucleotide triphosphates in the presence of appropriate DNA polymerases and/or with ligases.

As an alternative to such stepwise construction of a synthetic DNA chain, the cDNA corresponding to IGFBP-5 that was cloned to deduce the complete structure of IGFBP-5 is conveniently used. For example, an appropriate portion of the human IGFBP-5 DNA sequence encoding the 252-residue mature protein plus the signal sequence might be employed; the entire 1023 bp nucleic acid sequence is set forth as SEQ ID NO:5 in the appended SEQUENCE LISTING.

Alternatively, an appropriate portion of the rat IGFBP-5 DNA sequence encoding the 252-residue mature protein plus the signal sequence might be employed, with the entire 1630 bp nucleic acid sequence being set forth as SEQ ID NO:6 in the appended SEQUENCE LISTING.

In addition to the IGFBP-5-encoding sequence, a DNA chain should contain additional sequences depending upon vector construction considerations. Typically, a synthesized DNA chain has linkers at its ends to facilitate insertion into restriction sit s within a cl ning vector. A DNA chain may be constructed so as to

encod the IGFBP-5 amino acid sequences as a portion of a fusion polypeptide; and if so, it will generally contain terminal sequences that encode amino acid residue sequences that serve as proteolytic processing sites, whereby the IGFBP-5 polypeptide may be proteolytically cleaved from the remainder of the fusion polypeptide. The terminal portions of the synthetic DNA chain may also contain appropriate start and stop signals.

Accordingly, a double-stranded IGFBP-5-encoding DNA chain is constructed or modified with appropriate linkers for its insertion into a particular appropriate cloning vector. The cloning vector that is to be recombined to incorporate the DNA chain is selected appropriate to its viability and expression in a host organism or cell line, and the manner of insertion of the DNA chain depends upon factors particular to the host. For example, if the DNA chain is to be inserted into a vector for insertion into a prokaryotic cell, such as E. coli, the DNA chain will be inserted 3' of a promoter sequence, a Shine-Delgarno sequence (or ribosome binding site) that is within a 5' non-translated portion and an ATG start codon. The ATG start codon is appropriately spaced from the Shine-Delgarno sequence, and the encoding sequence is placed in correct reading frame with the ATG start codon. The cloning vector also provides a 3' non-translated region and a translation termination site. For insertion into a eukaryotic cell, such as a yeast cell or a cell line obtained from a higher animal, the IGFBP-5-encoding oligonucleotide sequence is appropriately spaced from a capping site and in correct reading frame with an ATG start signal. The cloning vector also provides a 3' non-translated region and a translation termination site.

Prokaryotic transformation vectors, such as pBR322, pMB9, Col E1, pCR1, RP4 and lambda-phag , are available for inserting a DNA chain of the length which

encodes IGFBP-5 with substantial assurance of at least some expression of the encoded polypeptide. Typically, such vectors are constructed or modified to have one or more unique restriction sites appropriately positioned relative to a promoter, such as the lac promoter. The DNA chain may be inserted with appropriate linkers into such a restriction site, with substantial assurance of production of IGFBP-5 in a prokaryotic cell line transformed with the recombinant vector. To assure proper reading frame, linkers of various lengths may be provided at the ends of the IGFBP-5-encoding sequences. Alternatively, cassettes, which include sequences, such as the 5' region of the lac Z gene (including the operator, promoter, transcription start site, Shine-Delgarno sequence and translation initiation signal), the regulatory region from the tryptophane gene (trp operator, promoter, ribosome binding site and translation initiator), and a fusion gene containing these two promoters called the trp-lac or commonly called the Tac promoter are available into which the synthetic DNA chain may be conveniently inserted and then the cassette inserted into a cloning vector of choice.

Similarly, eukaryotic transformation vectors, such as, the cloned bovine papilloma virus genome, the cloned genomes of the murine retroviruses, and eukaryotic cassettes, such as the pSV-2 gpt system (described by Mulligan and Berg, Nature 277, 108-114, 1979), the Okayama-Berg cloning system (Mol. Cell Biol. 2, 161-170, 1982), and the expression cloning vector recently described by Genetics Institute (Science 228, 810-815, 1985), are available which provide substantial assurance of at least some expression of IGFBP-5 in the transformed eukaryotic cell line.

As previously mentioned, a convenient way to ensure production of IGFBP-5 or a protein of a similar length is to produce the protein initially as a segment

of a gene-encoded fusion protein. In such case, the DNA chain is constructed so that the expressed protein has enzymatic processing sites flanking the IGFBP-5 amino acid residue sequences. An IGFBP-5-encoding DNA chain may be inserted, for example, into the beta-galactosidase gene for insertion into E. coli, in which case, the expressed fusion protein is subsequently cleaved with proteolytic enzymes to release the IGFBP-5 from beta-galactosidase peptide sequences. An advantage of inserting the IGFBP-5-encoding sequence so that the IGFBP-5 sequence is expressed as a cleavable segment of a fusion protein, e.g. as the IGFBP-5 sequence fused within the beta-galactosidase peptide sequence, is that the endogenous protein into which the IGFBP-5 sequence is inserted is generally rendered non-functional, thereby facilitating selection for vectors encoding the fusion protein.

The IGFBP-5 protein may also be reproduced in yeast using known recombinant DNA techniques. For example, plasmid pIGFBP-5, amplified in a pIGFBP-5-producing E. coli clone, is isolated and cleaved with Eco RI and Sal I. This digested plasmid is electrophoresed on an agarose gel allowing for the separation and recovery of the amplified pIGFBP-5 insert. The insert is inserted into the plasmid pYEp, a shuttle vector which can be used to transform both E. coli and Saccharomyces cerevisiae yeast. Insertion of the synthetic DNA chain at this point assures that the DNA sequence is under the control of a promoter, in proper reading frame from an ATG signal and properly spaced relative to a cap site. The shuttle vector is used to transform URA3, a strain of S. cerevisiae yeast from which the oratate monophosphate decarboxylase gene is deleted.

The transformed yeast is grown in medium to attain log growth. The yeast is separated from its

culture medium, and cell lysates are prepared. Pooled cell lysates are determined by RIA to be reactive with antibody raised against IGFBP-5, demonstrating that a protein containing IGFBP-5 protein segment is expressed within the yeast cells. The production of IGFBP-5 can be carried out in both prokaryotic and eukaryotic cell lines to provide protein for biological and therapeutic use. While IGFBP-5 synthesis is easily demonstrated using either bacteria or yeast cell lines, the synthetic genes should be insertable for expression in cells of higher animals, such as mammalian tumor cells. Such mammalian cells may be grown, for example, as peritoneal tumors in host animals, and IGFBP-5 harvested from the peritoneal fluid.

Although the above examples demonstrate that IGFBP-5 can be synthesized through recombinant DNA techniques, the examples do not purport to have maximized IGFBP-5 production. It is expected that subsequent selection of more efficient cloning vectors and host cell lines will increase the yield of IGFBP-5. Known gene amplification techniques for both eukaryotic and prokaryotic cells may be used to increase production of IGFBP-5. Secretion of the gene-encoded protein from the host cell line into the culture medium is also considered to be an important factor in obtaining synthetic IGFBP-5 in large quantities.

The availability of such mammalian IGFBP-5 proteins permit their use to complex and neutralize IGFs and these proteins should be useful in the treatment of conditions which are caused by an overabundance of IGFs, for example, certain types of breast or bone cancer. Administration of substantially pure monoclonal antibodies to IGFBP-5 have potential therapeutic applications to treat cases where it is desired to counteract the binding of IGFs, for example, in the modulation of bone growth.

Substantially pure IGFBP-5 protein can be routinely obtained having significantly higher purity than IGFBP-5 that is extracted from mammalian serum. IGFBP-5 proteins constitute only minor constituents of normal mammalian serum, being present in only very impure form, relative to other native proteins also present. Recombinant DNA techniques, for example, can be used to generate organisms or cell lines that produce the heterologous protein in significantly higher proportions relative to total protein, in the cellular material and/or the secretions thereof, than the proportions at which native IGFBP-5 are present. Because the starting material from which such synthetic IGFBP-5 proteins are isolated has a substantially greater concentration of the heterologous protein, purification techniques can fairly simply produce more highly purified IGFBP-5 fractions. Using appropriate isolation techniques, it is possible to routinely obtain IGFBP-5 proteins which are at least about 95% pure (by weight of total proteins) and which is herein referred to as substantially pure.

The protein should be administered under the guidance of a physician, and pharmaceutical compositions will usually contain the protein in conjunction with a conventional, pharmaceutically-acceptable carrier. For treatment, substantially pure synthetic IGFBP-5 or the nontoxic salts thereof, combined with a pharmaceutically acceptable carrier to form a pharmaceutical composition, are administered to mammals, including humans, either intravenously, subcutaneously, intramuscularly or orally. The required dosage will vary with the particular treatment and with the duration of desired treatment; however, it is anticipated that dosages between about 10 micrograms and about 1 milligram per kilogram of body weight per day will be used for therapeutic treatment, such as administration to persons afflicted with breast and/or bone cancers. Antibodies are administered in

proportionately appropriate amounts in accordance with known practices in this art.

Such protein may be administered in the form of pharmaceutically acceptable nontoxic salts, such as acid addition salts or metal complexes, e.g., with zinc, iron or the like (which are broadly considered as salts for purposes of this application). Illustrative of such acid addition salts are hydrochloride, hydrobromide, sulphate, phosphate, maleate, acetate, citrate, benzoate, succinate, malate, ascorbate, tartrate and the like. If the active ingredient is to be administered in tablet form, the tablet may contain a binder, such as tragacanth, corn starch or gelatin; a disintegrating agent, such as alginic acid; and a lubricant, such as magnesium stearate. If administration in liquid form is desired, sweetening and/or flavoring may be used, and intravenous administration in isotonic saline, phosphate buffer solutions or the like may be effected.

It may also be desirable to deliver IGFBP-5 over prolonged periods of time, for example, for periods of one week to one year from a single administration, and slow release, depot or implant dosage forms may be utilized. For example, a dosage form may contain a pharmaceutically acceptable non-toxic salt of the compound which has a low degree of solubility in body fluids, for example, an acid addition salt with the polybasic acid; a salt with a polyvalent metal cation; or combination of the two salts. A relatively insoluble salt may also be formulated in a gel, for example, an aluminum stearate gel. A suitable slow release depot formulation for injection may also contain IGFBP-5 or a salt thereof dispersed or encapsulated in a slow degrading, non-toxic or non-antigenic polymer such as a polylactic acid/polyglycolic acid polymer, for example, as described in U.S. Pat. No. 3,773,919. These compounds may also be formulated into silastic implants.

For purposes of this application, mammalian IGFBP-5 proteins should be considered to constitute proteins having the amino acid residue sequences set forth hereinbefore as well as naturally occurring amino acid sequence variants of other mammalian species and fragments of the foregoing having equivalent biological activity. Unless otherwise stated hereinbefore, all percentages are volume percents.

Although the invention has been described with regard to its preferred embodiments, which constitute the best mode presently known to the inventors, it should be understood that various changes and modifications as would be obvious to one having the ordinary skill in this art may be made without departing from the scope of the invention which is set forth in the claims appended hereto. For example, biologically active fragments of most proteins, shortened by the elimination of a sequence at the C-terminus or a sequence at the N-terminus or both, can be employed instead of the entire protein, and such fragments are considered to be equivalents of the mature protein IGFBP-5.

Particular features of the invention are emphasized in the claims which follow.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Shimasaki, Shunichi
Ling, Nicholas C.

(ii) TITLE OF INVENTION: Insulin-Like Growth Factor Binding Protein

(iii) NUMBER OF SEQUENCES: 10

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Fitch, Even, Tabin & Flannery
(B) STREET: 135 South LaSalle Street, Suite 900
(C) CITY: Chicago
(D) STATE: Illinois
(E) COUNTRY: United States
(F) ZIP: 60603-4277

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/658,410
(B) FILING DATE: 14-FEB-1991

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Watt, Phillip H.
(B) REGISTRATION NUMBER: 25,939
(C) REFERENCE/DOCKET NUMBER: 51145PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (312)372-7842
(B) TELEFAX: (312)372-7848

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Leu Gly Ser Phe Val His Xaa Glu Pro Xaa Asp Glu Lys Ala Leu Ser
1 5 10 15

Met Xaa Pro Pro Ser Pro Leu Gly Xaa Glu Leu Val Lys Glu Pro Gly
20 25 30

Xaa Gly Xaa Xaa
35

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ala Leu Ala Gly Xaa Pro Gly Xaa Gly Pro Gly Val Gln
1 5 10

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Val Ile Ser Val Val Leu Leu Leu Leu Ala Ala Cys Ala Val Pro
1 5 10 15

Ala Gln Gly Leu Gly Ser Phe Val His Cys Glu Pro Cys Asp Glu Lys
20 25 30

Ala Leu Ser Met Cys Pro Pro Ser Pro Leu Gly Cys Glu Leu Val Lys
35 40 45

Glu Pro Gly Cys Gly Cys Cys Met Thr Cys Ala Leu Ala Glu Gly Gln
50 55 60

Ser Cys Gly Val Tyr Thr Glu Arg Cys Ala Gln Gly Leu Arg Cys Leu
65 70 75 80

Pro Arg Gln Asp Glu Glu Lys Pro Leu His Ala Leu Leu His Gly Arg
                85   90   95

Gly Val Cys Leu Asn Glu Lys Ser Tyr Gly Glu Gln Thr Lys Ile Glu
            100   105   110

Arg Asp Ser Arg Glu His Glu Glu Pro Thr Thr Ser Glu Met Ala Glu
        115   120   125

Glu Thr Tyr Ser Pro Lys Val Phe Arg Pro Lys His Thr Arg Ile Ser
    130   135   140

Glu Leu Lys Ala Glu Ala Val Lys Lys Asp Arg Arg Lys Lys Leu Thr
145   150   155   160

Gln Ser Lys Phe Val Gly Gly Ala Glu Asn Thr Ala His Pro Arg Val
                165   170   175

Ile Pro Ala Pro Glu Met Arg Gln Glu Ser Asp Gln Gly Pro Cys Arg
            180   185   190

Arg His Met Glu Ala Ser Leu Gln Glu Phe Lys Ala Ser Pro Arg Met
        195   200   205

Val Pro Arg Ala Val Tyr Leu Pro Asn Cys Asp Arg Lys Gly Phe Tyr
    210   215   220

Lys Arg Lys Gln Cys Lys Pro Ser Arg Gly Arg Lys Arg Gly Ile Cys
225   230   235   240

Trp Cys Val Asp Lys Tyr Gly Met Lys Leu Pro Gly Met Glu Tyr Val
                245   250   255

Asp Gly Asp Phe Gln Cys His Ala Phe Asp Ser Ser Asn Val Glu
            260   265   270

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 272 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Val Leu Leu Thr Ala Val Leu Leu Leu Leu Ala Ala Tyr Ala Gly
1   5   10   15

Pro Ala Gln Ser Leu Gly Ser Phe Val His Cys Glu Pr Cys Asp Glu
            20   25   30

Lys Ala Leu Ser Met Cys Pro Pro Ser Pr Leu Gly Cys Glu Leu Val
35 40 45

Lys Glu Pro Gly Cys Gly Cys Cys Met Thr Cys Ala Leu Ala Glu Gly
50 55 60

Gln Ser Cys Gly Val Tyr Thr Glu Arg Cys Ala Gln Gly Leu Arg Cys
65 70 75 80

Leu Pro Arg Gln Asp Glu Glu Lys Pro Leu His Ala Leu Leu His Gly
85 90 95

Arg Gly Val Cys Leu Asn Glu Lys Ser Tyr Arg Glu Gln Val Lys Ile
100 105 110

Glu Arg Asp Ser Arg Glu His Glu Glu Pro Thr Thr Ser Glu Met Ala
115 120 125

Glu Glu Thr Tyr Ser Pro Lys Ile Phe Arg Pro Lys His Thr Arg Ile
130 135 140

Ser Glu Leu Lys Ala Glu Ala Val Lys Lys Asp Arg Arg Lys Lys Leu
145 150 155 160

Thr Gln Ser Lys Phe Val Gly Gly Ala Glu Asn Thr Ala His Pro Arg
165 170 175

Ile Ile Ser Ala Pro Glu Met Arg Gln Glu Ser Glu Gln Gly Pro Cys
180 185 190

Arg Arg His Met Glu Ala Ser Leu Gln Glu Leu Lys Ala Ser Pro Arg
195 200 205

Met Val Pro Arg Ala Val Tyr Leu Pro Asn Cys Asp Arg Lys Gly Phe
210 215 220

Tyr Lys Arg Lys Gln Cys Lys Pro Ser Arg Gly Arg Lys Arg Gly Ile
225 230 235 240

Cys Trp Cys Val Asp Lys Tyr Gly Met Lys Leu Pro Gly Met Glu Tyr
245 250 255

Val Asp Gly Asp Phe Gln Cys His Thr Phe Asp Ser Ser Asn Val Glu
260 265 270

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1023 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 57..872

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 117..872

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
CCCTGCACTC TCGCTCTCCT GCCCCACCCC GAGGTAAAGG GGGCGACTAA GAGAAG            56

ATG GTG TTG CTC ACC GCG GTC CTC CTG CTG CTG GCC GCC TAT GCG GGG         104
Met Val Leu Leu Thr Ala Val Leu Leu Leu Leu Ala Ala Tyr Ala Gly
-20                 -15                 -10                  -5

CCG GCC CAG AGC CTG GGC TCC TTC GTG CAC TGC GAG CCC TGC GAC GAG         152
Pro Ala Gln Ser Leu Gly Ser Phe Val His Cys Glu Pro Cys Asp Glu
                  1               5                  10

AAA GCC CTC TCC ATG TGC CCC CCC AGC CCC CTG GGC TGC GAG CTG GTC         200
Lys Ala Leu Ser Met Cys Pro Pro Ser Pro Leu Gly Cys Glu Leu Val
        15                  20                  25

AAG GAG CCG GGC TGC GGC TGC TGC ATG ACC TGC GCC CTG GCC GAG GGG         248
Lys Glu Pro Gly Cys Gly Cys Cys Met Thr Cys Ala Leu Ala Glu Gly
    30                  35                  40

CAG TCG TGC GGC GTC TAC ACC GAG CGC TGC GCC CAG GGG CTG CGC TGC         296
Gln Ser Cys Gly Val Tyr Thr Glu Arg Cys Ala Gln Gly Leu Arg Cys
45                  50                  55                  60

CTC CCC CGG CAG GAC GAG GAG AAG CCG CTG CAC GCC CTG CTG CAC GGC         344
Leu Pro Arg Gln Asp Glu Glu Lys Pro Leu His Ala Leu Leu His Gly
                65                  70                  75

CGC GGG GTT TGC CTC AAC GAA AAG AGC TAC CGC GAG CAA GTC AAG ATC         392
Arg Gly Val Cys Leu Asn Glu Lys Ser Tyr Arg Glu Gln Val Lys Ile
            80                  85                  90

GAG AGA GAC TCC CGT GAG CAC GAG GAG CCC ACC ACC TCT GAG ATG GCC         440
Glu Arg Asp Ser Arg Glu His Glu Glu Pro Thr Thr Ser Glu M t Ala
        95                  100                 105
```

```
GAG GAG ACC TAC TCC CCC AAG ATC TTC CGG CCC AAA CAC ACC CGC ATC     488
Glu Glu Thr Tyr Ser Pro Lys Ile Phe Arg Pro Lys His Thr Arg Ile
    110                 115                 120

TCC GAG CTG AAG GCT GAA GCA GTG AAG AAG GAC CGC AGA AAG AAG CTG     536
Ser Glu Leu Lys Ala Glu Ala Val Lys Lys Asp Arg Arg Lys Lys Leu
125                 130                 135                 140

ACC CAG TCC AAG TTT GTC GGG GGA GCC GAG AAC ACT GCC CAC CCC CGG     584
Thr Gln Ser Lys Phe Val Gly Gly Ala Glu Asn Thr Ala His Pro Arg
                145                 150                 155

ATC ATC TCT GCA CCT GAG ATG AGA CAG GAG TCT GAG CAG GGC CCC TGC     632
Ile Ile Ser Ala Pro Glu Met Arg Gln Glu Ser Glu Gln Gly Pro Cys
            160                 165                 170

CGC AGA CAC ATG GAG GCT TCC CTG CAG GAG CTC AAA GCC AGC CCA CGC     680
Arg Arg His Met Glu Ala Ser Leu Gln Glu Leu Lys Ala Ser Pro Arg
        175                 180                 185

ATG GTG CCC CGT GCT GTG TAC CTG CCC AAT TGT GAC CGC AAA GGA TTC     728
Met Val Pro Arg Ala Val Tyr Leu Pro Asn Cys Asp Arg Lys Gly Phe
    190                 195                 200

TAC AAG AGA AAG CAG TGC AAA CCT TCC CGT GGC CGC AAG CGT GGC ATC     776
Tyr Lys Arg Lys Gln Cys Lys Pro Ser Arg Gly Arg Lys Arg Gly Ile
205                 210                 215                 220

TGC TGG TGC GTG GAC AAG TAC GGG ATG AAG CTG CCA GGC ATG GAG TAC     824
Cys Trp Cys Val Asp Lys Tyr Gly Met Lys Leu Pro Gly Met Glu Tyr
                225                 230                 235

GTT GAC GGG GAC TTT CAG TGC CAC ACC TTC GAC AGC AGC AAC GTT GAG     872
Val Asp Gly Asp Phe Gln Cys His Thr Phe Asp Ser Ser Asn Val Glu
            240                 245                 250

TGATGCGTCC CCCCCCAACC TTTCCCTCAC CCCCTCCCAC CCCCAGCCCC GACTCCAGCC     932

AGCGCCTCCC TCCACCCCAG GACGCCACTC ATTTCATCTC ATTTAAGGGA AAAATATATA     992

TCTATCTATT TGAAAAAAAA AAAAAAAACC C                                  1023
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1630 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 551..1363

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 608..1363

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
CTTTTTTTTC CTCTCCCTTG ATTTCAACAT TTTCCCGATC TTCTTGGCGC CGCCGACGCC      60

TCTTACCTGT TCTGCGGCAGC GCGGCGAGCTG GCAGCTGAGA GAAGTGGGGG TGCGTTTTAG      120

GTTTTAAGCA AAGGCAAAAA AAAAAATTAA GCCAAATCCA TTTTTTTTCC TTCACCTCTC      180

CCCGTTTCAA GGCCTCCAAG ATCATTATTT CTGTAGCTTT GGGGTGAGCG ATTCTGTGTT      240

TCTCTTCATC ACCCCTCCAA TTCTGCCCCG ATCCCGCCTG GGTCTCCACT CACTGCGTGC      300

ACCTGGCGCG CCTCTTTTTT TTTTCACCCC CAACCTGTTG CAAGTCTTTA ATCCTTGCAA      360

TTGGGACTTG CGTGCAGGCA CCTGAATCCT CCTTGCCTCA TATTTTGCAA GTGTTTGGAG      420

GACAGCACCT GCTTTACCTG CAAGAGATAT TTTTTAAAAA AAAAAAAATC TCCAGGCTCC      480

CTCTTGGCCC CTTTCTCCAC ACACTCTCGC TCTCCTGCCC CGCCCCGAGG TAAAGCCAGA      540

CTCGGAAAAA ATG GTG ATC AGC GTG GTC CTC CTG CTG CTG GCC GCC TGT      589
           Met Val Ile Ser Val Val Leu Leu Leu Leu Ala Ala Cys
           -19             -15                 -10

GCC GTG CCG GCT CAA GGC CTG GGC TCT TTC GTG CAT TGT GAA CCC TGC      637
Ala Val Pro Ala Gln Gly Leu Gly Ser Phe Val His Cys Glu Pro Cys
    -5                  1               5                  10

GAC GAG AAA GCT CTG TCC ATG TGT CCC CCC AGC CCT CTG GGC TGT GAG      685
Asp Glu Lys Ala Leu Ser Met Cys Pro Pro Ser Pro Leu Gly Cys Glu
                15                  20                  25

CTG GTC AAA GAG CCC GGC TGT GGC TGC TGC ATG ACT TGC GCC CTG GCG      733
Leu Val Lys Glu Pro Gly Cys Gly Cys Cys Met Thr Cys Ala Leu Ala
            30                  35                  40
```

```
GAG GGA CAG TCG TGT GGC GTC TAC ACT GAG CGC TGC GCC CAG GGT TTG      781
Glu Gly Gln Ser Cys Gly Val Tyr Thr Glu Arg Cys Ala Gln Gly Leu
        45                  50                  55

CGC TGT CTC CCC CGG CAG GAT GAG GAG AAG CCG CTG CAC GCC CTG CTG      829
Arg Cys Leu Pro Arg Gln Asp Glu Glu Lys Pro Leu His Ala Leu Leu
    60                  65                  70

CAC GGC CGC GGG GTT TGC CTC AAC GAA AAG AGC TAC GGC GAG CAA ACC      877
His Gly Arg Gly Val Cys Leu Asn Glu Lys Ser Tyr Gly Glu Gln Thr
75                  80                  85                  90

AAG ATA GAG AGA GAC TCT CGG GAG CAT GAG GAA CCC ACC ACC TCC GAG      925
Lys Ile Glu Arg Asp Ser Arg Glu His Glu Glu Pro Thr Thr Ser Glu
                95                  100                 105

ATG GCT GAG GAG ACC TAC TCC CCG AAG GTC TTC CGG CCC AAG CAC ACT      973
Met Ala Glu Glu Thr Tyr Ser Pro Lys Val Phe Arg Pro Lys His Thr
            110                 115                 120

CGC ATT TCC GAG CTG AAG GCC GAG GCT GTG AAG AAG GAT CGC AGA AAG     1021
Arg Ile Ser Glu Leu Lys Ala Glu Ala Val Lys Lys Asp Arg Arg Lys
        125                 130                 135

AAG CTG ACC CAG TCT AAG TTT GTG GGG GGC GCG GAG AAC ACT GCC CAC     1069
Lys Leu Thr Gln Ser Lys Phe Val Gly Gly Ala Glu Asn Thr Ala His
    140                 145                 150

CCC CGA GTC ATC CCT GCA CCT GAG ATG AGA CAG GAA TCT GAC CAA GGC     1117
Pro Arg Val Ile Pro Ala Pro Glu Met Arg Gln Glu Ser Asp Gln Gly
155                 160                 165                 170

CCC TGC CGC AGA CAC ATG GAA GCT TCC CTC CAG GAG TTC AAA GCC AGC     1165
Pro Cys Arg Arg His Met Glu Ala Ser Leu Gln Glu Phe Lys Ala Ser
                175                 180                 185

CCA CGC ATG GTG CCC CGT GCC GTG TAC CTG CCC AAC TGT GAC CGC AAA     1213
Pro Arg Met Val Pro Arg Ala Val Tyr Leu Pro Asn Cys Asp Arg Lys
            190                 195                 200

GGA TTC TAC AAG AGA AAG CAG TGC AAG CCT TCT CGT GGC CGC AAA CGT     1261
Gly Phe Tyr Lys Arg Lys Gln Cys Lys Pro Ser Arg Gly Arg Lys Arg
        205                 210                 215

GGC ATC TGC TGG TGT GTG GAC AAG TAT GGG ATG AAG CTG CCG GGC ATG     1309
Gly Ile Cys Trp Cys Val Asp Lys Tyr Gly Met Lys Leu Pro Gly Met
    220                 225                 230

GAG TAC GTC GAT GGG GAC TTT CAG TGC CAC GCC TTC GAC AGC AGT AAC     1357
Glu Tyr Val Asp Gly Asp Phe Gln Cys His Ala Phe Asp Ser Ser Asn
235                 240                 245                 250
```

GTT GAG TGACGCGTCC CCTCCCTTCC TCCCCCTTTC CTACCCCCCA GCCCCAACTC 1413
Val Glu

CAGCCAGCGC CTCCCTCCAC CCCAGGACGT CACTCATTTC ATCTCATTTA GGGGAAATAT 1473

ATATACATAT ATATTTGAGG AAACTGAGGA CCTCGGAATC TCTAGCAAGG GCTAAGGAGA 1533

CACTCCCCAT TCCCGACCCC GGAAACGTAT TCCTATTTGA AGCAAGTTGA ACGGACAGAG 1593

AAGGGAAGAA GAGAAGGGGC AAGAAGGAGC GAGGAAT 1630

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ala Gln Xaa Pro Gly Xaa Gly Gln Gly Val Gln Thr Gly Xaa Pro Gly
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ala Gly Pro Cys Trp Cys Val Asp Ser Arg Pro Asn Pro Gly Gly Val
1 5 10 15

Gln Asp Thr Glu Met Gly Pro Cys Arg
20 25

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Leu Ala Gln Cys Pro Gly Cys Gly Gln Gly Val Gln Thr Gly Cys Pro
1   5   10   15

Gly Gly Cys Ala Glu Glu Glu Asp Gly Gly Xaa Pro Arg Glu Arg Val
            20   25   30

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Glu Gly Gln Gln Cys Gly Val Tyr Thr Pro Asn Cys Ala Pro Gly Leu
1   5   10   15

Gln Cys Gln Pro Pro Glu Glu Asp Gln Ala Pro Leu Arg
            20   25

WHAT IS CLAIMED IS:

1. Recombinant DNA consisting essentially of a DNA sequence encoding an IGFBP-5 having the amino acid sequence SEQ ID NO: 4 or a homologous naturally occurring variant DNA sequence of another animal species, or encoding an N-terminally shortened fragment of such an IGFBP-5.

2. The DNA of Claim 1 which encodes the IGFBP-5 protein of SEQ ID NO:4 less the 20 residues at the N-terminus.

3. The DNA of Claim 1 having all or a substantial portion of either the nucleotide sequence SEQ ID NO: 5 or of a homologous naturally occurring variant nucleotide sequence of another mammalian species.

4. DNA according to Claim 3 wherein said homologous variant sequence is the nucleotide sequence SEQ ID NO:6.

5. A replicable recombinant DNA expression vector which includes the DNA of Claim 1, said vector being capable of expressing the DNA in a microorganism or cell culture wherein said vector is inserted.

6. The vector of Claim 5 wherein, upon expression, the protein of the amino acid sequence SEQ ID NO:4 less the 20 residues at the N-terminus is produced.

7. Recombinant host cells transformed with the vector of Claim 6.

8. A microorganism transformed with the vector of Claim 6, said microorganism being capable of expressing the DNA encoding the IGFBP-5.

9. A cell culture capable of expressing DNA encoding an IGFBP-5 protein, which cell culture is obtained by transforming a cell line with the vector of Claim 5.

10. A protein which has SEQ ID NO: 4 or a highly homologous, naturally occurring, variant protein of another mammalian species, or a C-terminally shortened fragment of either.

11. The protein according to Claim 10 which has SEQ ID NO: 4 less the 20 residues at the N-terminus.

12. Antibodies which bind to and inactivate the protein according to Claim 11.

13. The protein according to Claim 10 which has SEQ ID NO:3.

14. The protein according to Claim 10 which has the SEQ ID NO:3 less the 19 residues at the N-terminus.

15. Antibodies which bind to and inactivate the protein according to Claim 14.

16. A method of treating a patient afflicted with breast and/or bone cancer, which method comprises administering an effective amount of the protein in accordance with Claim 10.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US92/01196

I. CLASSIFICATI N OF SUBJECT MATTER (if several classification symbols apply, indicate all)[3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC (5): C12P 33/02; C07K 3/00; A61K 37/24
US CL : 435/69.1; 530/350, 399

II. FIELDS SEARCHED

Minimum Documentation Searched[4]

Classification System	Classification Symbols
U.S.	435/69.1; 530/350, 399

Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched[5]

dialog, aps, intelligenetics

III. DOCUMENTS CONSIDERED TO BE RELEVANT[14]

Category*	Citation of Document,[16] with indication, where appropriate, of the relevant passages[17]	Relevant to Claim No.[18]
Y, P	J. Biol. Chem. Vol. 266, No. 16, issued 05 June 1991, S. Shimasaki et al., "Identification of five different Insulin-like growth factor binding proteins (IGFBPs) from adult rat serum and molecular cloning of a novel IGFBP-5 in rat and human", pages 10646-10653, see entire document.	1 - 16
Y, P	J. Biol. Chem. Vol. 266, No. 15, issued 15 May 1991, M.C. Kiefer et al., "Identification and molecular cloning of two new 30-kD insulin-like growth factor binding proteins isolated from adult human serum", pages 9043-9049, see entire document.	1 - 16
Y, P	Biochem. Biophys. Res. Commun., Vol. 176, No. 1 issued 15 April 1991, M.C. Kiefer et al., "Molecular cloning of a new human insulin-like growth factor binding protein", pages 219-225, see entire document.	1 - 16

* Special categories of cited documents:[15]

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search[2]	Date of Mailing of this International Search Report[2]
14 MAY 1992	22 MAY 1992
International Searching Authority[1] ISA/US	Signature of Authorized Officer[20] Deborah E. Freese for Karen Cochrane Carlson, Ph.D.

Form PCT/ISA/210 (second sheet)(May 1986) 8